

# On the Convexity of Image of a Multidimensional Quadratic Map

Anatoly Dymarsky <sup>a,1</sup>

<sup>a</sup>Center for Theoretical Physics, MIT, Cambridge, MA, USA, 02139

<sup>1</sup>Skolkovo Institute of Science and Technology,  
Novaya St. 100, Skolkovo, Moscow Region, Russia, 143025

## Abstract

We study convexity of the image of a general multidimensional quadratic map. We split the full image into two parts by an appropriate hyperplane such that one part is compact and formulate a sufficient condition for its convexity. We propose a simple way to identify such convex parts of the full image which can be used in practical applications. By shifting the hyperplane to infinity we extend the sufficient condition for the convexity to the full image of the quadratic map. We also discuss the connection of our findings with the classical question of convexity of the joint numerical range of  $m$ -tuple of hermitian matrices. In particular we outline novel sufficient conditions for the joint numerical range to be convex.

**Keywords:** convexity, quadratic transformation, multidimensional quadratic map, vector-valued quadratic forms, joint numerical range, Polyak convexity principle

## 1 Introduction and Main Results

In this paper we consider a multidimensional quadratic map of general form  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  (or  $f : \mathbb{C}^n \rightarrow \mathbb{R}^m$ ) defined by an  $m$ -tuple of symmetric (hermitian) matrices  $A_i$ , an  $m$ -tuple of vectors  $v_i \in \mathbb{R}^n$  (or  $v_i \in \mathbb{C}^n$ ), and a vector  $f^0 \in \mathbb{R}^m$ ,<sup>2</sup>

$$f_i(x) = x^* A_i x - v_i^* x - x^* v_i + f_i^0. \quad (1.1)$$

---

<sup>1</sup>Permanent address.

<sup>2</sup>Symbol \* stands for transposition or hermitian conjugation depending on the context.

An important question arising in many applications is when the image of the quadratic map (1.1),

$$\mathcal{F}(f) = \{f(x) : x \in \mathbb{V}\} \subseteq \mathbb{R}^m, \quad \mathbb{V} = \mathbb{R}^n \text{ or } \mathbb{V} = \mathbb{C}^n, \quad (1.2)$$

is convex. This question was previously studied in case of small  $m = 2, 3$  or when only a few of  $A_i$ 's are linearly independent, see [1, 2, 3] and [4] for references and a brief historic overview. Identifying necessary and sufficient conditions for the convexity of (1.2) for general  $m$  remains an open problem which we investigate in this paper.

Although the results concerning convexity of  $\mathcal{F}(f)$  as a whole are scarce, something can be said about convexity of the image of  $f$  locally. In particular, for any non-linear map  $f$  there is an upper bound on the size of a ball  $B_\varepsilon(x_0) = \{x : |x - x_0|^2 \leq \varepsilon^2\} \subset \mathbb{V}$  such that its image  $f(B_\varepsilon(x_0)) = \{f(x) : x \in B_\varepsilon(x_0)\}$  is convex [5]. Because of a very broad scope of this result the corresponding bound on  $\varepsilon$  is always finite and  $f(B_\varepsilon(x_0))$  can never fully cover  $\mathcal{F}(f)$ . The bound of [5] was improved in [6] for the case of a general quadratic map (1.1) and an ellipsoid-shaped ball  $B_\varepsilon^+(x_0) = \{x : |x - x_0|_+^2 \leq \varepsilon^2\}$  defined through a positive-definite matrix  $A_+$ ,  $|x|_+^2 := x^* A_+ x$ ,

$$\varepsilon_{\max}^2(x_0, A_+) := \lim_{\varepsilon \rightarrow 0^+} \min_{c \in \mathcal{C}} |(c \cdot A - \lambda_{\min}^+(c \cdot A) A_+ + \epsilon)^{-1} c \cdot (v - A x_0)|_+^2, \quad (1.3)$$

$$\mathcal{C} = \{c : c \in \mathbb{R}^m, |c|^2 = 1, \lambda_{\min}^+(c \cdot A) \leq 0\}. \quad (1.4)$$

Above we introduced  $\lambda_{\min}^+(A)$  to denote the smallest generalized eigenvalue of  $A$  with respect to  $A_+$ .<sup>3</sup> In (1.3) and in what follows a sum of a matrix and a number understood in a sense that the number is multiplied by the identity matrix of an appropriate size. The dot product  $\cdot$  stands for the standard scalar product in the Euclidean space. Also notice the appearance of the the “plus” norm in (1.3).

Reference [6] proves that  $f(B_\varepsilon^+(x_0))$  is strictly convex for  $\varepsilon \leq \varepsilon_{\max}$ . Now, if for any non sign-definite combination  $c \cdot A \not\equiv 0$ ,  $c \in \mathbb{R}^m \setminus \{0\}$ , the projection of the vector  $c \cdot (v - A x_0)$  on the eigenspace corresponding to the smallest generalized eigenvalue of  $c \cdot A$  is non-vanishing,  $\varepsilon_{\max}$  is infinite and the image  $f(B_\varepsilon^+(x_0))$  is strictly convex for any  $\varepsilon$ . Consequently we have the following proposition.

**Proposition 1.** If for any point  $x_0 \in \mathbb{V}$  and any  $A_+ \succ 0$  the limit  $\varepsilon_{\max}(x_0, A_+)$  is infinite the full image  $\mathcal{F}(f)$  is convex. This is a novel sufficient condition for the convexity of image of a general multidimensional quadratic map.

---

<sup>3</sup>The generalized eigenvalues are defined through  $(A - \lambda^+ A_+)x = 0$  for some non-trivial  $x$ .

The proof is trivial. For any two points  $y_1, y_2 \in \mathcal{F}(f)$  we consider their pre-images  $x_1, x_2 \in \mathbb{V}$  (any pre-images if there are many). For a significantly large  $\varepsilon$ ,  $x_1, x_2 \in B_\varepsilon^+(x_0)$  and consequently all points  $y(t) = y_1(1-t) + y_2t$  for  $1 \geq t \geq 0$  lie within  $f(B_\varepsilon^+(x_0)) \subset \mathcal{F}(f)$ .

What happens if  $\varepsilon_{\max}$  is very large but not infinite? Strictly speaking there is not much we can say about the convexity of  $\mathcal{F}(f)$  in this case. It is reasonable to expect that  $\mathcal{F}(f)$  might still be very close to be convex because the image of a very large ball  $B_\varepsilon^+(x_0) \subset \mathbb{V}$  is convex. Unfortunately some points  $x \in \mathbb{V}$ ,  $x \notin B_\varepsilon^+(x_0)$  might be mapped in  $\mathbb{R}^m$  finite distance away from the origin for any, even infinitely large  $\varepsilon$  and this could spoil convexity of the full image. If somehow we could arrange for  $\|y=f(x)\|$ ,  $x \in \mathbb{V}$ ,  $x \notin B_\varepsilon^+(x_0)$ , to be large with some appropriate norm  $\|\cdot\|$  when  $\varepsilon$  is large than not only  $\mathcal{F}(f)$  would be convex when  $\varepsilon_{\max}^2$  is infinite, but also we could outline a compact part of  $\mathcal{F}(f)$  (an intersection of  $\mathcal{F}(f)$  with a “ball” of a certain size defined by the norm  $\|\cdot\|$ ), which is convex when  $\varepsilon_{\max}^2$  is finite. This idea, which we develop in section 2, leads to the following result.

For any vector  $c_+ \in \mathbb{R}^m \setminus \{0\}$  such that the combination  $A_+ := c_+ \cdot A$  is positive-definite,  $A_+ \succ 0$ , let us define the point  $x_0 = A_+^{-1}(c_+ \cdot v)$ . This is the unique point where the supporting hyperplane orthogonal to  $c_+$  “touches”  $\mathcal{F}(f)$ . Let us introduce the following limit

$$z_{\max} := \lim_{\epsilon \rightarrow 0^+} \min_{c \in \mathcal{C}} |(c \cdot A - \lambda_{\min}^+(c \cdot A)A_+ + \epsilon)^{-1} c \cdot (v - A x_0)|_+^2, \quad (1.5)$$

$$\mathcal{C} = \{c : c \in \mathbb{R}^m, |c|^2 = 1, c \cdot c_+ = 0\}. \quad (1.6)$$

**Proposition 2.** The compact set  $\mathcal{F}(f, c_+, z_{\max}) = \{f(x) : x \in \mathbb{V}, |x - x_0|_+^2 \leq z_{\max}\} \subset \mathcal{F}(f)$  is convex. If  $z_{\max}$  is infinite the whole image  $\mathcal{F}(f)$  is convex.

*Comment 1.* The convex set  $\mathcal{F}(f, c_+, z_{\max})$  can be defined as a compact part of  $\mathcal{F}(f)$  lying in a half-space defined by a certain hyperplane  $\mathcal{H}_{c_+}(c_+ \cdot f(x_0) + z_{\max})$ , where for any  $c \in \mathbb{R}^m \setminus \{0\}$  the corresponding orthogonal hyperplane is defined as

$$\mathcal{H}_c(F) = \{y : y \in \mathbb{R}^m, c \cdot y = F\} \subset \mathbb{R}^m. \quad (1.7)$$

The hyperplane  $\mathcal{H}_{c_+}(c_+ \cdot f(x_0))$  is the supporting hyperplane of  $\mathcal{F}(f)$  perpendicular to  $c_+$ . Hence  $\mathcal{F}(f, c_+, z_{\max})$  is the part of  $\mathcal{F}(f)$  bounded by two parallel hyperplanes separated by the distance  $z_{\max}$  apart. This is illustrated in Fig. 1b.

*Comment 2.* The two criteria for convexity (Proposition 1 and Proposition 2) are

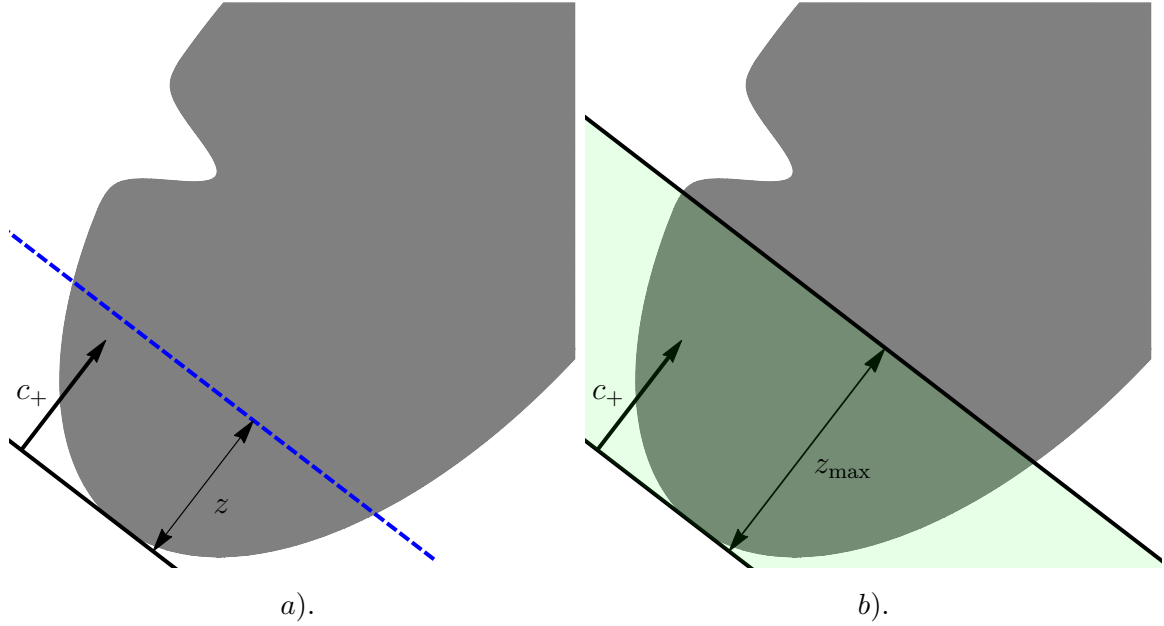


Figure 1: (a) An illustration of the main idea. A hyperplane orthogonal to  $c_+$  touches  $\mathcal{F}(f)$  only at one point. If the boundary is smooth and strongly convex around that point a parallel hyperplane located a short distance apart will carve a compact subregion of  $\mathcal{F}(f)$ . (b)  $z_{\max}$  defines a maximal subregion of  $\mathcal{F}(f)$  which is stably convex.

complimentary to each other in the following sense. In (1.3)  $x_0$  must be a regular point of  $f$  lest  $\varepsilon_{\max}(x_0)$  vanishes. On the contrary in (1.5)  $f(x_0)$  belongs to the boundary  $\partial\mathcal{F}(f)$ .

*Comment 3.* The sufficient condition of Proposition 1 depends on the choice of  $x_0, A_+$ . If  $\varepsilon_{\max}(x_0, A_+)$  is infinite for some  $x_0, A_+$  it might be finite for some other  $x_0, A_+$ . On the contrary if  $z_{\max}$  (1.5) is infinite for some  $c_+$  (and corresponding  $x_0, A_+$ ), it will be infinite for all other choices of  $c_+$  as well. If  $z_{\max}$  is finite its value depends on  $c_+$ . This is explained in section 2.1.

We derive the expression for  $z_{\max}$  (1.5) and prove Proposition 2 in section 2, while Proposition 1 directly follows from the results of [6]. In section 3 we discuss how our results pertaining to convexity of  $\mathcal{F}(f)$  can be related to the classical question of convexity of the joint numerical range of  $m$ -tuple of symmetric (hermitian) matrices. In particular we formulate novel sufficient conditions for convexity of the joint numerical range in section 3.1. Finally, in section 4 we discuss different ways of calculating  $z_{\max}$ . While a precise but computationally intense method is discussed in section 4,

in section 4.1 we propose an easy-to-calculate conservative estimate of  $z_{\max}$ . In this way we formulate a practical way to outline a convex part of  $\mathcal{F}(f)$ . This method can be used in applications to solve various problems of optimization and control.

## 2 Convexity of $\mathcal{F}(f)$

Let us first approach the question of convexity of  $\mathcal{F}(f)$  locally. From now on we assume that the  $m$ -tuple of matrices  $A_i$  is *definite* in the sense of reference [8], i.e. there is a positive-definite combination  $A_+ := c_+ \cdot A \succ 0$  for some  $c_+ \in \mathbb{R}^m \setminus \{0\}$ . The corresponding supporting hyperplane  $\mathcal{H}_{c_+}(z_0)$  touches  $\mathcal{F}(f)$  only at one point  $y_0 = f(x_0)$ ,

$$x_0 = A_+^{-1} v_+ , \quad z_0 = -v_+^* A_+^{-1} v_+ , \quad v_+ := c_+ \cdot v . \quad (2.1)$$

This is schematically depicted in Fig. 1a. Provided the boundary of  $\mathcal{F}(f)$  is smooth and strongly convex at  $y_0$  it is tempting to say that the compact part of  $\mathcal{F}(f)$  lying in a half-space defined by the hyperplane  $\mathcal{H}_{c_+}(z_0 + z)$ ,  $z > 0$ , would be convex, at least for very small  $z$ . It is easy to see that the compact part of  $\mathcal{F}(f)$  bounded by  $\mathcal{H}_{c_+}(z_0 + z)$ ,  $z > 0$ , is the image of the ball  $B_z^+(x_0) = \{x : |x - x_0|_+^2 \leq z\}$  under the map  $f$ . We would like to find an upper bound on  $z$  such that the image  $f(B_z^+(x_0))$  is convex. This problem almost identically repeats the question studied by Polyak [5] for general  $f$ , and further investigated in [6] in case of quadratic  $f$ , with one crucial distinction: point  $x_0$  is not a regular point of  $f$ .

The following observation drastically simplifies further analysis: each ellipsoid  $|x - x_0|_+^2 = z$  is mapped into its own hyperplane  $\mathcal{H}_{c_+}(z_0 + z)$ . Hence the convexity of  $f(B_z^+(x_0))$  requires convexity of  $f(|x|_+^2 = z')$  for all  $z' \leq z$ . Up to a translation and a trivial change of basis in  $\mathbb{V}$  the image  $f(|x - x_0|_+^2 = z)$  is nothing but the inhomogeneous joint numerical range, the notion we introduced in [6],

$$\mathcal{F}(\mathbf{A}, \mathbf{v}) = \{y_i : \exists x, y_i = x^* \mathbf{A}_i x - \mathbf{v}_i^* x - x^* \mathbf{v}_i, |x|^2 = 1\} , \quad (2.2)$$

for an  $m$ -tuple of symmetric (hermitian) matrices  $\mathbf{A}_i$  and a  $m$ -tuple of vectors  $\mathbf{v}_i$ . There it was proven that so far

$$\lim_{\epsilon \rightarrow 0^+} \min_{c \neq 0} |(c \cdot \mathbf{A} - \lambda_{\min}(c \cdot \mathbf{A}) + \epsilon)^{-1} c \cdot \mathbf{v}| \geq 1 , \quad (2.3)$$

$\mathcal{F}(\mathbf{A}, \mathbf{v})$  is strictly convex (strongly convex for strong inequality) and smooth. In fact (2.3) is a criterion for stable convexity, i.e. impossibility to ruin convexity of  $\mathcal{F}(\mathbf{A}, \mathbf{v})$  by an infinitesimal deformation of  $\mathbf{A}, \mathbf{v}$ .

Applying this directly to the image  $f(|x - x_0|_+^2 = z)$  will fail because the latter is “flat”, i.e. it lies within the hyperplane  $\mathcal{H}_{c_+}(z_0 + z)$  and hence can not be strictly convex. This is easy to fix by considering an orthogonal projection  $\varphi$  of  $\mathbb{R}^m$  on  $\mathcal{H}_{c_+} \simeq \mathbb{R}^{m-1}$ . Now the criterion (2.3) can be applied directly to the  $(m - 1)$ -dimensional inhomogeneous joint numerical range  $(\varphi \circ f)(|x - x_0|_+^2 = z)$ ,

$$z_{\max} := \lim_{\epsilon \rightarrow 0^+} \min_{c \neq 0} \left| (c \cdot A - \lambda_{\min}^+(c \cdot A)A_+ + \epsilon)^{-1} c \cdot (v - Ax_0) \right|_+^2 \geq z. \quad (2.4)$$

The minimum (2.4) is taken over the space of the equivalence classes  $c \in \mathbb{R}^m$ ,  $c \simeq c + \mu c_+$ ,  $\forall \mu \in \mathbb{R}$ . Since the minimized expression is homogeneous in  $c$  the condition  $c \neq 0$  could be substituted by  $|c|^2 = 1$ ,  $c \cdot c_+ = 0$ . (Notice, that if  $c \propto c_+$ , (2.4) vanishes.)

Clearly, convexity of  $\mathcal{F}(f) \cap \mathcal{H}_{c_+}(z_0 + z)$  for all  $z \leq z_{\max}$  is a necessary condition for the convexity of  $f(B_{z_{\max}}^+(x_0))$  but may not be sufficient. To establish convexity of  $f(B_{z_{\max}}^+(x_0))$  let us choose an arbitrary vector  $c \in \mathbb{R}^m \setminus \{0\}$  and find an intersection of  $f(B_{z_{\max}}^+(x_0))$  with a supporting hyperplane  $\mathcal{H}_c(F_c)$  orthogonal to  $c$  which touches  $f(B_{z_{\max}}^+(x_0))$  “from below”,

$$F_c = \min_{x \in B_{z_{\max}}^+(x_0)} c \cdot f(x). \quad (2.5)$$

For  $c$  collinear with  $c_+$  the answer is simple. When  $c \cdot c_+ > 0$  the hyperplane  $\mathcal{H}_{c_+}(z_0)$  intersects  $f(B_{z_{\max}}^+(x_0))$  at the unique point  $y_0$ . When  $c \cdot c_+ < 0$  the intersection  $\mathcal{H}_{c_+}(z_0 + z_{\max}) \cap f(B_{z_{\max}}^+(x_0))$  is a  $(m - 1)$ -dimensional convex set  $f(|x - x_0|_+^2 = z_{\max})$ . For  $c$  not collinear with  $c_+$  we can first find a conditional minimum for  $|x - x_0|_+^2 = z$  and then minimize with respect to  $z$  in the interval  $[0, z_{\max}]$ . This problem was solved in [6] where it was shown that the minimum is achieved at such  $z$  that  $\lambda(z)$ , uniquely determined by the conditions

$$\left| (c \cdot A - \lambda(z)A_+)^{-1} c \cdot (v - Ax_0) \right|_+^2 = z, \quad \lambda(z) \leq \lambda_{\min}^+(c \cdot A), \quad (2.6)$$

is equal to  $\min\{0, \lambda(z_{\max})\}$ . Since  $\lambda(z)$  is a monotonically increasing function of  $z$  on the interval  $[0, z_{\max}]$  such a point is unique. Correspondingly the supporting hyperplane  $\mathcal{H}_c(F_c)$  intersects  $f(B_{z_{\max}}^+(x_0))$  at a unique point. Hence  $\partial \text{Conv}[f(B_{z_{\max}}^+(x_0))] \subset$

$f(B_{z_{\max}}^+(x_0))$  where  $\text{Conv}$  stands for convex hull. Now, to prove that  $f(B_{z_{\max}}^+(x_0))$  is convex we need to show that any point of  $\text{Conv}[f(B_{z_{\max}}^+(x_0))]$  belongs to  $f(B_{z_{\max}}^+(x_0))$ . This is obviously true because an intersection of  $f(B_{z_{\max}}^+(x_0))$  with the hyperplane  $\mathcal{H}_{c_+}(F)$  for any  $F$  is either empty or convex. This finishes the proof of Proposition 2.

## 2.1 Geometrical meaning of $z_{\max}$

To interpret  $z_{\max}$  geometrically we would need to understand different scenarios of how  $\mathcal{F}(f)$  may intersect with its supporting hyperplanes. Let us consider a vector  $c \in \mathbb{R}^m \setminus \{0\}$  and find a supporting hyperplane to  $\mathcal{F}(f)$  that is orthogonal to  $c$ . There are several possible scenarios. First,  $c \cdot A$  is sign-definite. The corresponding supporting hyperplane intersects  $\mathcal{F}(f)$  at the unique point  $f(x)$ ,  $x = (c \cdot A)^{-1}c \cdot v$ . Second,  $c \cdot A$  has both positive and negative eigenvalues. There is no corresponding supporting hyperplane in this case because  $\mathcal{F}(f)$  stretches to infinity in both directions along  $c$ . Finally,  $c \cdot A$  is semi-definite and degenerate. There are two possibilities (Fredholm alternative) in this case. If the equation  $(c \cdot A)x = c \cdot J$  admits no solution, there is no supporting hyperplane to  $\mathcal{F}(f)$  orthogonal to  $c$  because  $\mathcal{F}(f)$  stretches to infinity in both directions along  $c$ . Another option is when there is a whole linear space of solutions of  $(c \cdot A)x = c \cdot J$ . Each such solution  $x$  corresponds to a point  $f(x)$  from the boundary  $\partial\mathcal{F}(f)$  belonging to the same supporting hyperplane orthogonal to  $c$ .

In case a supporting hyperplane intersects  $\mathcal{F}(f)$  over more than one point we would like to call all points of  $\partial\mathcal{F}(f)$  belonging to this supporting hyperplane a “flat edge”.<sup>4</sup> If  $\partial\mathcal{F}(f)$  includes “flat edges”,  $\mathcal{F}(f)$  can not be strictly convex and certainly  $\mathcal{F}(f)$  can not be stably convex. In general such  $\mathcal{F}(f)$  will not be convex at all.

Now let us look at the definition of  $z_{\max}$  (1.5). For any  $c$ ,  $c \cdot c_+ = 0$ , the limit  $\epsilon \rightarrow 0^+$  in (1.5) will be finite only if the equation

$$(\mathbf{c} \cdot A)(x + x_0) = \mathbf{c} \cdot v, \quad \mathbf{c}(c) = c - \lambda_{\min}^+(c \cdot A)c_+, \quad (2.7)$$

has nontrivial solution(s). Since  $\mathbf{c} \cdot A$  is degenerate and positive semi-definite, existence of nontrivial solutions is the same as the existence of a “flat edge” orthogonal to  $\mathbf{c}$ .

---

<sup>4</sup>More precisely we should define “flat edge” as an intersection of  $\mathcal{F}(f)$  with a supporting hyperplane if the pre-image of this intersection consists of more than one point. But a “flat edge” defined this way can consist of one point only if all matrices  $A_i$  have a common zero eigenvector which is in contradiction with the assumption that the set of  $A_i$ ’s is definite.

Hence  $z_{\max}$  will be infinite unless there is a supporting hyperplane touching  $\mathcal{F}(f)$  at more than one point. The latter is the property of  $f$  and  $\mathcal{F}(f)$  and does not depend on the choice of  $c_+$  and  $x_0, A_+$ .

It is easy to show that for any vector  $c$  such that  $\mathbf{c}(c)$  is orthogonal to a “flat edge” (namely,  $(\mathbf{c} \cdot A) \succeq 0$ ,  $(\mathbf{c} \cdot A) \not\equiv 0$ , the space of solutions  $\mathcal{S} = \{x : (\mathbf{c} \cdot A)x = \mathbf{c} \cdot v\}$  is non-trivial), the limit  $\epsilon \rightarrow 0^+$  in (1.5) calculates the minimum value of  $c_+ \cdot f(x)$ ,  $x \in \mathcal{S}$ . Consequently  $z_{\max}$  is the distance from the supporting hyperplane orthogonal to  $c_+$  to a closest point  $y \in \partial\mathcal{F}(f)$  belonging to any “flat edge” inside  $\partial\mathcal{F}(f)$ . This is depicted in Fig. 2a. This also explains that  $\mathcal{F}(f, c_+, z)$  is stably convex for  $z \leq z_{\max}$  and is not stably convex  $z > z_{\max}$ .

A comment is in order. If  $\mathcal{F}$  were compact, absence of “flat edges” would immediately guarantee that the outer boundary  $\partial\mathcal{F}$  is a boundary of the convex hull of  $\mathcal{F}$  while the latter is strictly convex. But in a non-compact case this is not true. Fig. 2b provides an non-compact example without “flat edges” when the outer boundary is not the boundary of the convex hull. Hence the proof that the absence of “flat edges” ( $z_{\max} \rightarrow \infty$ ) implies that the outer boundary  $\partial\mathcal{F}$  confines a convex set presented in section 2 was not superfluous.

### 3 Connection with the joint numerical range

In this section we would like to look at the convexity of  $\mathcal{F}(f)$  from a slightly different angle. Shifting  $f_i^0$  by a constant does not affect convexity and therefore without loss of generality we assume  $f_i^0 = 0$ . Then  $\mathcal{F}(f)$  can be thought of as an intersection of the image of the “extended” *homogeneous* quadratic map  $\mathbf{f} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{m+1}$  (or  $\mathbf{f} : \mathbb{C}^{n+1} \rightarrow \mathbb{R}^{m+1}$ ) and a hyperplane

$$\mathcal{F}(f) = \mathcal{F}(\mathbf{f}) \cap \mathcal{H}_{e_{m+1}}(1) , \quad (3.1)$$

where  $e_{m+1}$  is the  $(m+1)$ -th basis vector and

$$\mathbf{f}_I = \mathbf{x}^* \mathbf{A}_I \mathbf{x} , \quad I = 1, \dots, m+1 , \quad \mathbf{x} \in \mathbb{R}^{n+1} \text{ (or } \mathbb{C}^{n+1}) , \quad (3.2)$$

$$\mathbf{A}_i = \left( \begin{array}{c|c} A_i & -v_i \\ \hline -v_i^* & 0 \end{array} \right) , \quad i = 1, \dots, m , \quad \mathbf{A}_{m+1} = \left( \begin{array}{c|c} 0_{n \times n} & 0_{n \times 1} \\ \hline 0_{1 \times n} & 1 \end{array} \right) , \quad (3.3)$$

$$\mathcal{H}_{\mathbf{c}}(F) = \{y : y \in \mathbb{R}^{m+1} , \mathbf{c} \cdot y = F\} \subset \mathbb{R}^{m+1} , \quad \forall \mathbf{c} \in \mathbb{R}^{m+1} . \quad (3.4)$$



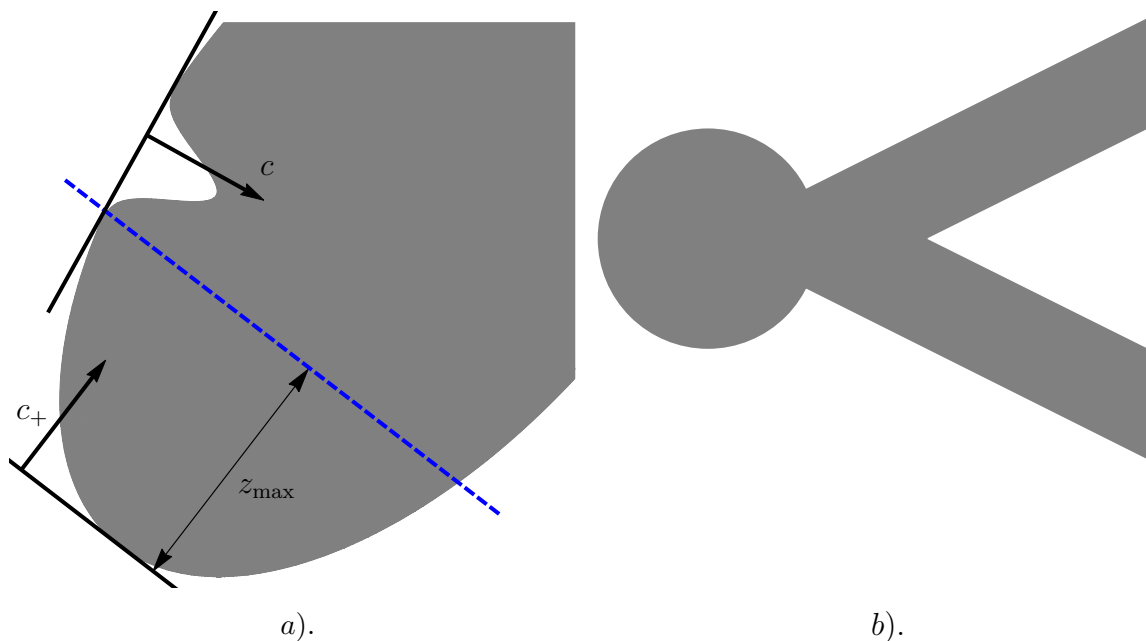


Figure 2: (a) Geometrical meaning of  $z_{\max}$ . It is the distance from a supporting hyperplane orthogonal to  $c_+$  to the closest point belonging to a “flat edge”. (b) A non-convex figure which intersects any supporting hyperplane at exactly one point. (The shevron-shape strips continue to infinity.)

Obviously, convexity of  $\mathcal{F}(\mathbf{f})$  would imply convexity of  $\mathcal{F}(f)$ . Validity of the converse statement is a subject of the following discussion. On general grounds conic structure of  $\mathcal{F}(\mathbf{f})$  and the relation (3.1) do not imply convexity of  $\mathcal{F}(\mathbf{f})$  but only of  $\mathcal{F}(\mathbf{f}) \setminus \mathcal{H}_{e_{m+1}}(0)$ .

Possible advantage of introducing  $\mathcal{F}(\mathbf{f})$  is that convexity of homogeneous quadratic maps was studied previously in [8]. There is a general relation between definite homogeneous quadratic maps and joint numerical ranges of some auxiliary matrices. Geometrically  $\mathcal{F}(\mathbf{f})$  is a cone. If  $\mathbf{f}$  is definite, i.e. there is a linear combination  $c_+ \cdot \mathbf{A} \succ 0$ , the base of the cone  $\mathcal{F}(\mathbf{f}) \cap \mathcal{H}_{c_+}(F)$ ,  $F > 0$ , is compact and equal (up to a linear isomorphism) to a joint numerical range  $\mathcal{F}(\mathcal{A})$  of  $m$  matrices  $\mathcal{A}_i$  which can be explicitly constructed from  $\mathbf{A}_I$ . Hence proving convexity of  $\mathcal{F}(\mathbf{f})$  is the same as establishing convexity of  $\mathcal{F}(\mathcal{A})$ . The latter is a question with a long and rich history.

For an  $m$ -tuple of symmetric or hermitian matrices  $\mathcal{A}_i$  the joint numerical range is defined as follows

$$\mathcal{F}(\mathcal{A}) = \{y_i : \exists x, x \in \mathbb{V}, |x|^2 = 1, y_i = x^* \mathcal{A}_i x\} \subset \mathbb{R}^m, \quad (3.5)$$

where  $\mathbb{V} = \mathbb{R}^n$  or  $\mathbb{V} = \mathbb{C}^n$  correspondingly. The question of convexity of  $\mathcal{F}$  goes back

to Housdorff and Toeplitz [9] who proved that  $\mathcal{F}$  is always convex for  $m = 2$  hermitian matrices (i.e.  $\mathbb{V} = \mathbb{C}^n$ ) and  $n > 1$ . There are numerous results for small  $m = 2, 3$  [10, 11, 12, 13] (also see [14] for references) and a few specific conditions rendering  $\mathcal{F}$  non-convex. The case of general  $m, n$  is not fully understood, although there is a sufficient condition that guarantees that  $\mathcal{F}(\mathcal{A})$  is strongly convex and smooth: if for all linear combinations of  $\mathcal{A}_i$  the dimension of the eigenspace corresponding to the lowest eigenvalue is the same [7]. In terms of the corresponding map  $f$  this is the condition that for any linear combination  $\mathbf{c} \cdot \mathbf{A} \succeq 0$ ,  $\mathbf{c} \cdot \mathbf{A} \not\equiv 0$ , dimension of  $\ker(\mathbf{c} \cdot \mathbf{A})$  is the same. This condition is a generalization of *roundness* defined in [8] which guarantees “roundness” (strict convexity and smoothness) of the base of the cone  $\mathcal{F}(f)$ .

We will see later that in our case of interest (3.3) the condition that  $\dim(\ker(\mathbf{c} \cdot \mathbf{A}))$  remains the same for different  $\mathbf{c}$  is not satisfied. Thus the results of [7], [8] do not help to establish convexity of  $\mathcal{F}(f)$ . In the following we will reverse the logic and extend the sufficient condition for convexity of  $\mathcal{F}(f)$ , Proposition 2, to  $\mathcal{F}(f)$  and  $\mathcal{F}(\mathcal{A})$ . In this way we formulate new criteria for the convexity of the joint numerical range.

From now on we assume that  $z_{\max}$  given by (1.5) is infinite, which implicitly assumes the set of  $A_i$ ’s is definite. Let us show that the set of  $\mathbf{A}_I$ ’s given by (3.3) is definite as well. Starting from an appropriate  $c_+$ , let us consider the vector  $\mathbf{c}_+ \in \mathbb{R}^{m+1}$ ,

$$\mathbf{c}_{+i} = c_{+i}, \quad i = 1, \dots, m, \quad \mathbf{c}_{+(m+1)} > v_+^* A_+^{-1} v_+ . \quad (3.6)$$

Then the matrix  $\mathbf{A}_+ := \mathbf{c}_+ \cdot \mathbf{A}$  is positive-definite as follows from the Sylvester’s criterion.

Infinite value of  $z_{\max}$  implies that for any  $c$ ,  $c \cdot A \succeq 0$ ,  $c \cdot A \not\equiv 0$ , equation  $(c \cdot A)x = c \cdot v$  has no solution (see section 2.1). This immediately implies that any  $\mathbf{x} \in \ker(\mathbf{c} \cdot \mathbf{A})$ , where  $\mathbf{c} \cdot \mathbf{A} \succeq 0$  is either trivial or must satisfy  $\mathbf{x}^* \mathbf{A}_{m+1} \mathbf{x} > 0$ , unless  $\mathbf{c} \propto e_{m+1}$ . Hence  $\dim(\ker(\mathbf{c} \cdot \mathbf{A})) = 1$  for all appropriate  $\mathbf{c}$ , except for the special case  $\dim(\ker(\mathbf{A}_{m+1})) = n$ . Presence of this exceptional direction makes it impossible to apply the results of [7, 8]. Indeed these works focused on strictly convex and smooth  $\mathcal{F}(A)$ , properties guaranteed by constancy of  $\dim(\ker(\mathbf{c} \cdot \mathbf{A})) = 1$  for all appropriate  $\mathbf{c}$ ,  $\mathbf{c} \cdot \mathbf{A} \succeq 0$ ,  $\mathbf{c} \cdot \mathbf{A} \not\equiv 0$ . But in our case  $\mathcal{F}(A)$  has a “flat edge” perpendicular to the direction of  $e_{m+1}$  (after projection on  $\mathcal{H}_{c_+}(1)$ ), which is a direct consequence of  $\dim(\ker(\mathbf{c} \cdot \mathbf{A})) = n$  for one particular direction  $\mathbf{c} \propto e_{m+1}$ . The joint numerical range of this kind is shown in Fig. 3.

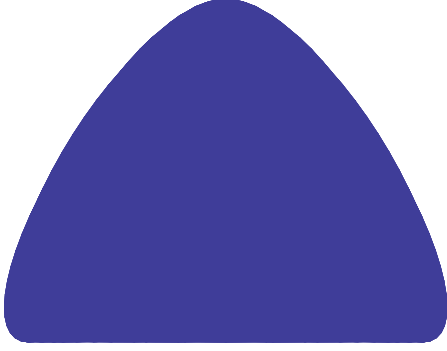


Figure 3: Join numerical range for two matrices

$$\mathcal{A}_1 = \begin{pmatrix} 2 & 0 & -\frac{3}{2} \\ 0 & -2 & \frac{3}{2} \\ -\frac{3}{2} & \frac{3}{2} & 0 \end{pmatrix}, \quad \mathcal{A}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let us establish convexity of the “flat edge” of  $\mathcal{F}(\mathcal{A})$  which is isomorphic to  $\mathfrak{F}(0)$  where we introduce

$$\mathfrak{F}(F) = \mathcal{F}(\mathbf{f}) \cap \mathcal{H}_{\mathbf{c}_+}(1) \cap \mathcal{H}_{e_{m+1}}(F), \quad F \geq 0. \quad (3.7)$$

For any  $F > 0$ ,  $\mathfrak{F}(F)$  is isomorphic to  $\mathcal{F}(f) \cap \mathcal{H}_{\mathbf{c}_+}(z_0 + z) = f(|x - x_0|_+^2 = z)$  where corresponding  $\mathbf{c}_+$  (and hence  $x_0, z_0$ ) is related to  $\mathbf{c}_+$  through (3.6) and  $z$  is some function of  $F$  and other parameters. The limit  $F \rightarrow 0$  corresponds  $z \rightarrow \infty$ . For any sufficiently small  $F_0 > F > 0$ , where  $z(F_0) = z_0$ , we proved that  $\mathfrak{F}(F) = f(|x - x_0|_+^2 = z)$  is strongly convex. Hence by continuity  $\mathfrak{F}(0)$  is convex as well.

The rest of the proof closely follows the logic outlined in section 2. Let us consider a vector  $\mathbf{c} \in \mathbb{R}^{m+1}$ ,  $|\mathbf{c}|^2 = 1$ ,  $\mathbf{c} \cdot \mathbf{c}_+ = 0$ , and find an intersection of  $\mathcal{F}(\mathcal{A}) \simeq \mathcal{F}(\mathbf{f}) \cap \mathcal{H}_{\mathbf{c}_+}(1)$  with the supporting hyperplane  $\mathcal{H}_{\mathbf{c}}(F_c)$  orthogonal to  $\mathbf{c}$  touching  $\mathcal{F}(\mathcal{A})$  “from below”

$$F_c = \min_{y \in \mathcal{F}(\mathbf{f}) \cap \mathcal{H}_{\mathbf{c}_+}(1)} \mathbf{c} \cdot y. \quad (3.8)$$

For any  $\mathbf{c}$  the intersection consists of a unique point except for  $\mathbf{c} = e_{m+1}$  when it is a convex “flat edge”. Hence the outer boundary  $\partial\mathcal{F}(\mathcal{A})$  is the boundary of the convex hull of  $\mathcal{F}(\mathcal{A})$ ,  $\partial\text{Conv}[\mathcal{F}(\mathcal{A})] \subset \mathcal{F}(\mathcal{A})$ . Finally, since the intersection  $\mathcal{F}(\mathbf{f}) \cap \mathcal{H}_{\mathbf{c}_+}(1) \cap \mathcal{H}_{e_{m+1}}(F)$  is convex (or empty) for any  $F$  we conclude that all points confined by  $\partial\mathcal{F}(\mathcal{A})$  belong to  $\mathcal{F}(\mathcal{A})$ . This establishes convexity of  $\mathcal{F}(\mathbf{f})$  and  $\mathcal{F}(\mathcal{A})$  provided the criterion for convexity of  $\mathcal{F}(f)$ , Proposition 2, is satisfied.

### 3.1 New Criteria for Convexity for the Joint Numerical Range

The proof of the convexity of  $\mathcal{F}(\mathcal{A})$  can be cast in a form of a self-contained criterion for convexity based only on the properties of matrices  $\mathcal{A}_i$ .

Let us consider a joint numerical range (3.5) defined by an  $m$ -tuple of  $n \times n$  symmetric (hermitian) matrices  $\mathcal{A}_i$ . If for any linear combination  $c \cdot \mathcal{A}$ ,  $c \in \mathbb{R}^m \setminus \{0\}$  its smallest eigenvalue is not degenerate,  $\mathcal{F}(\mathcal{A})$  is convex [7]. If for any linear combination  $c \cdot \mathcal{A}$ ,  $c \in \mathbb{R}^m \setminus \{c : c = \mu e, \mu \geq 0\}$ , where  $e \in \mathbb{R}^m \setminus \{0\}$  is a fixed vector, its smallest eigenvalue is not degenerate, but the smallest eigenvalue of  $e \cdot \mathcal{A}$  is  $(n - 1)$ -times degenerate,  $\mathcal{F}(\mathcal{A})$  is convex if the following condition is satisfied. We add the identity matrix  $\mathbb{I}_{n \times n}$  to the set of  $\mathcal{A}_i$ 's thus bringing the total number of matrices to  $(m + 1)$ . By changing a basis in  $\mathbb{V}$  we bring  $e \cdot \mathcal{A}$  to the form of  $\mathbf{A}_{m+1}$  from (3.3). By taking a linear combination of  $e \cdot \mathcal{A}$  with other  $m$  matrices we bring them to the form of  $\mathbf{A}_i$  from (3.3) and in this way define  $m$  vectors  $v_i$  and  $(n - 1) \times (n - 1)$  matrices  $A_i$ . Now, if the auxiliary map  $f$  defined by  $A_i, v_i$  through (1.1) satisfies the convexity criteria of Proposition 2,  $(z_{\max} \rightarrow \infty)$ , then  $\mathcal{F}(\mathcal{A})$  is convex.

In fact there is another sufficient criterion of convexity of  $\mathcal{F}(\mathcal{A})$  “buried” inside the proof in section 3. Indeed, the “flat edge”  $\mathfrak{F}(0)$  is the joint numerical range associated with the homogeneous quadratic map  $y_i = x^* A_i x$ . Hence the second sufficient condition for the convexity of the joint numerical range can be formulated as follows.

Let us consider a joint numerical range (3.5) defined by an  $m$ -tuple of  $n \times n$  symmetric (hermitian) matrices  $\mathcal{A}_i$ . By adding the identity matrix to the set of  $\mathcal{A}_i$ 's we define an  $(m + 1)$ -tuple of matrices  $\{A_I\} = \{\mathcal{A}_i\} \cup \{\mathbb{I}_{n \times n}\}$ . If for any  $(m + 1)$ -tuple of vectors  $v_I$  the corresponding quadratic map

$$f_I = x^* A_I x - v_I^* x - x^* v_I, \quad (3.9)$$

satisfies the convexity criteria of Proposition 2,  $(z_{\max} \rightarrow \infty)$ , then  $\mathcal{F}(\mathcal{A})$  is convex.

The two criteria for convexity formulated above lack in elegance. The need to introduce the auxiliary map  $f$  suggests there might be a better way to establish convexity of  $\mathcal{F}(\mathcal{A})$ . We leave the task of formulating the criteria for convexity free of any reference to  $f$  and  $A_i, v_i$  for the future.

## 4 Different Approaches to Calculating $z_{\max}$

In section 2 we proved that a compact part of  $\mathcal{F}(f)$  lying in a half-space defined by the hyperplane  $\mathcal{H}_{c+}(z_0 + z_{\max})$  is compact. To make this a practical method of carving a compact subregion within  $\mathcal{F}(f)$  suitable for applications we would need to

be able to determine  $z_{\max}$  for a given  $c_+$ . A straightforward approach would be to use the definition (1.5). We prefer to rewrite  $c \cdot A - \lambda_{\min}^+(c \cdot A)A_+$  as  $\mathbf{c} \cdot A$ , where  $\mathbf{c}(c)$  is given by (2.7). Notice that matrix  $\mathbf{c} \cdot A$  for all  $c \in \mathbb{R}^m$ ,  $c \cdot c_+ = 0$ , exhausts all positive-semidefinite combinations of  $A_i$  with a non-trivial kernel. Hence  $c$  orthogonal to  $c_+$  parametrize the boundary of the convex cone  $\mathcal{K}^+$  of the positive-semidefinite linear combinations of  $A_i$ . Furthermore  $c \cdot (v - Ax_0) = \mathbf{c} \cdot (v - Ax_0)$  and therefore minimization problem (1.5) can be naturally defined on the boundary  $\partial\mathcal{K}^+$ ,

$$z_{\max} = \lim_{\epsilon \rightarrow 0^+} \min_{c \in \partial\mathcal{K}^+} |(c \cdot A + \epsilon)^{-1} c \cdot (v - Ax_0)|_+^2, \quad (4.1)$$

$$\mathcal{K}^+ = \{c : c \in \mathbb{R}^m, c \cdot A \succeq 0\}. \quad (4.2)$$

Unfortunately this problem is not convex. Moreover extending this problem to the interior of  $\mathcal{K}^+$  is not a viable option because  $\lim_{\epsilon \rightarrow 0^+} \min_{c \in \mathcal{K}^+} |(c \cdot A + \epsilon)^{-1} c \cdot (v - Ax_0)|_+^2 = 0$  even when  $z_{\max} > 0$ . We conclude that the minimization problem (1.5) can not be immediately reduced to a convex optimization problem admitting an efficient solution.

This motivates us to approach the problem of calculating  $z_{\max}$  from a slightly different angle. As was discussed in section 2.1 the minimized expression in (4.1) is finite in the limit  $\epsilon \rightarrow 0$  only when there are nontrivial solutions of  $\mathbf{c}(c) \cdot Ax = \mathbf{c}(c) \cdot (v - Ax_0)$ . In turn this implies one can define a matrix  $A(c) = \mathbf{c} \cdot A$ ,  $\mathbf{c}_{+i} = \mathbf{c}_{+i}$ ,  $i = 1, \dots, m$ , and fine-tune  $\mathbf{c}_{+(m+1)}$  such that  $A(c)$  is positive semi-definite and admits at least two zero eigenvectors. Hence a better strategy to find  $z_{\max}$  might be to first identify all  $\mathbf{c} \in \mathbb{R}^m$  such that  $\mathbf{c} \cdot A \succeq 0$ ,  $\text{rank}(\mathbf{c} \cdot A) \leq n - 2$ ,  $\mathbf{c}^T = \{\mathbf{c}^T, 1\}^T$ , and calculate  $z_{\mathbf{c}} = |(\mathbf{c} \cdot A)^{-1} \mathbf{c} \cdot (v - Ax_0)|_+^2$ . The value of  $z_{\mathbf{c}}$  is the minimum  $\min_{x \in \mathcal{S}} |x - x_0|_+^2$  where  $\mathcal{S} = \{x : \mathbf{c} \cdot Ax = \mathbf{c} \cdot v\}$ . To find  $z_{\max}$  one would need to find a minimal  $z_{\mathbf{c}}$  for all such  $\mathbf{c}$ .

The condition  $\text{rank}(\mathbf{c} \cdot A) = n - 2$  leads to the following two algebraic conditions: the determinant of  $\mathbf{c} \cdot A$  is zero as well as the sum of all diagonal first minors. These two algebraic conditions, together with the condition  $\mathbf{c} \in \mathbb{R}^m$  and the set of algebraic inequalities resulting from  $\mathbf{c} \cdot A \succeq 0$  (Sylvester's criterion) determine the set of vectors  $\mathbb{C}$ .

Because of the algebraic nature of the aforementioned conditions it might be useful to relax the condition that the components of  $\mathbf{c}$  are real, solve the remaining constraints and then impose reality of  $\mathbf{c}_i$  in the end. Once  $\mathbf{c}_i$  are complex, matrix  $\mathbf{c} \cdot A$  may not be diagonalizable, hence requiring that two lowest terms of its characteristic polynomial vanish may not be enough. Instead we could require that all  $n^2$  of its first

minors vanish. Although not all of these necessary conditions would be independent, this approach might be preferable if solving the corresponding constraints with help of a computer algebra system. By further imposing  $c_i \in \mathbb{R}$  for all  $i$  and  $c \cdot A \succeq 0$  we can determine the  $m - 4$  dimensional family of vectors  $c$ .

## 4.1 Conservative Estimate of $z_{\max}$

Calculating  $z_{\max}$  exactly could be a difficult task which may require solving a non-convex minimization problem or finding roots of a large system of algebraic equations. Nevertheless for many practical application it would be enough to have an easy-to-calculate conservative estimate  $z_{\text{est}} \leq z_{\max}$ . A very similar problem of estimating  $\varepsilon_{\max}^2$  (1.3) was addressed in [6] and here we employ the same strategy. The first step is the inequality

$$z_{\text{est}} := \min_{c \in \mathcal{C}} \frac{|c \cdot \tilde{v}|^2}{\|c \cdot \tilde{A} - \lambda_{\min}(c \cdot \tilde{A})\|^2} \leq z_{\max} , \quad (4.3)$$

$$\mathcal{C} = \{c : c \in \mathbb{R}^m, |c|^2 = 1, c \cdot c_+ = 0\} , \quad (4.4)$$

where  $\tilde{v}_i = \lambda(v_i - A_i x_0)$ ,  $\tilde{A}_i = \lambda A_i \lambda^*$  and  $A_+ = (\lambda^* \lambda)^{-1}$ . Because (4.3) is homogeneous in  $c$  the condition  $|c|^2 = 1$  can be substituted by  $c^* g c = 1$  for some  $g$  which is positive-definite on the orthogonal complement to  $c_+$  within  $\mathbb{R}^m$ . Let's choose  $g_{ij} = \text{Re}(\tilde{v}_i^* \tilde{v}_j)$  which satisfies this requirement (obviously  $g_{ij}$  is non-negative; if it develops a zero eigenvalue on the orthogonal complement to  $c_+$ , then, alas,  $z_{\max}$  would be zero anyway). It is convenient to diagonalize  $g$  and bring it to the standard form  $\Lambda^{-1} g \Lambda = \text{diag}(1, \dots, 1, 0)$  with the help of some non-generate real-valued  $m \times m$  matrix  $\Lambda$ . Now we can define  $\hat{A}_i = \Lambda_i^j \tilde{A}_j$  and since  $\Lambda_m^i \sim c_+^i$  we introduce a new  $(m - 1)$ -dimensional vector  $\hat{c}$  with the components  $\hat{c}_1, \dots, \hat{c}_{m-1}$ ,

$$z_{\text{est}} = \left( \max_{|\hat{c}|^2=1} \left( \lambda_{\max}(\hat{c} \cdot \hat{A}) - \lambda_{\min}(\hat{c} \cdot \hat{A}) \right)^2 \right)^{-2}, \quad \hat{c} \in \mathbb{R}^{m-1} . \quad (4.5)$$

Furthermore using the inequality  $\lambda_{\max}(A) - \lambda_{\min}(A) \leq 2 \max\{\lambda_{\max}(A), \lambda_{\max}(-A)\}$  we arrive at the following approximate conservative estimate of  $z_{\text{est}}$ ,

$$\left( \max_{|\hat{c}|^2=1} 2 \lambda_{\max}(\hat{c} \cdot \hat{A}) \right)^{-2} \leq z_{\text{est}} . \quad (4.6)$$

Next step would be to calculate or estimate  $\max_{|\hat{c}|^2=1} \lambda_{\max}(\hat{c} \cdot \hat{A})$  – the problem which we addressed in [6].

In fact the estimate (4.6) can be improved if we notice that  $\lambda_{\max}(A) - \lambda_{\min}(A)$  is invariant under the shifts of  $A$  by the identity matrix  $A \rightarrow A + \mu \mathbb{I}_{n \times n}$ . Hence the better estimate for  $z_{\text{est}}$  would be

$$z = 4^{-1} \left( \min_{\mu} \max_{|\hat{c}|^2=1} \left( \lambda_{\max}(\hat{c} \cdot \hat{A}) + \mu(\hat{c}) \right) \right)^{-2} \leq z_{\text{est}} , \quad (4.7)$$

where the minimum is taken over the space of functions  $\mu(\hat{c})$  satisfying  $\mu(-\hat{c}) = -\mu(\hat{c})$ . Obviously we can restrict the class of functions  $\mu$  by paying the price of somewhat deteriorating the quality of the estimate. For example  $\mu$  could be chosen to be a linear function  $\mu(\hat{c}) = \hat{c} \cdot \mu$  defined by a vector  $\mu_i$ . Depending on the chosen method to estimate  $\max_{|\hat{c}|^2=1} \lambda_{\max}(\hat{c} \cdot (\hat{A} + \mu))$  we can either find a minimum with respect to  $\mu_i$  analytically or leave it to numerical analysis.

In case minimization with respect to  $\mu_i$  is too difficult to handle one guideline to choose  $\mu_i$  could be the following. The inequality  $\lambda_{\max}(A) - \lambda_{\min}(A) \leq 2 \max\{\lambda_{\max}(A), \lambda_{\max}(-A)\}$  is saturated when  $\lambda_{\max}(A) = -\lambda_{\min}(A)$  and therefore it makes sense to choose  $\mu_i$  such that any combination  $\hat{c} \cdot (\hat{A} + \mu)$  has both positive and negative eigenvalues, for example  $\mu_i = -\text{Tr}(\hat{A}_i)$ .

Thus, for one of the approximate methods for calculating  $\max_{|\hat{c}|^2=1} \lambda_{\max}(\hat{c} \cdot (\hat{A} + \mu))$  proposed in [6] the choice  $\mu_i = -\text{Tr}(\hat{A}_i)$  is in fact the result of the exact minimization with respect to  $\mu_i$ , yielding

$$z_{\text{est}} \geq z \geq 2^{-2} \lambda_{\max}^{-2}(M) , \quad M_{ij} = \text{Tr}(\hat{A}_i \hat{A}_j) - \text{Tr}(\hat{A}_i) \text{Tr}(\hat{A}_j) / m . \quad (4.8)$$

## Acknowledgments

I would like to thank Vasily Pestun and Konstantin Turitsyn for useful discussions and gratefully acknowledge support from the grant RFBR 12-01-00482.

## References

- [1] B. T. Polyak, “Convexity of quadratic transformations and its use in control and optimization,” *Journal of Optimization Theory and Applications*, 99, 553-583 (1998).

- [2] A. Beck, "On the convexity of a class of quadratic mappings and its application to the problem of finding the smallest ball enclosing a given intersection of balls," *Journal of Global Optimization*, Vol. 39, Issue 1, September 2007, p. 113 - 126.
- [3] Y. Xia, "On Local Convexity of Quadratic Transformations," [arXiv:1405.6042 [math.OC]].
- [4] I. Pólik, T. Terlaky, "A Survey of S-lemma," *SIAM review*, 49(3), 371-418 (2007).
- [5] B. T. Polyak, "Convexity of Nonlinear Image of a Small Ball with Applications to Optimization," *Set-Valued Analysis*, **9**(1/2), 159-168 (2001).  
 B. T. Polyak, "Local Programming," *Comp. Math. and Math. Phys.*, **41**(9), 1259-1266 (2001).  
 B. T. Polyak, "The convexity principle and its applications," *Bull.Braz.Math.Soc. (N.S.)*, **34**(1), 59-75 (2003).
- [6] A. Dymarsky, "Convexity of a Small Ball Under Quadratic Map," [arXiv:1410.1553 [math.OC]].
- [7] E. Gutkin, E. A. Jonckheere, M. Karow, "Convexity of the joint numerical range: topological and differential geometric viewpoints," *Linear Algebra Appl.* 376 (2004), 143171.
- [8] J. Sheriff, "The Convexity of Quadratic Maps and the Controllability of Coupled Systems," Ph.D. thesis, Harvard University, 2013, UMI publication number 3567063.
- [9] O. Toeplitz, "Das algebraische Analogon zu einem Satze von Fejér," *Mathematische Zeitschrift*, 1918, Vol. 2, Issue 1-2, pp. 187-197.  
 F. Hausdorff, "Der Wertevorrat einer Bilinearform," *Mathematische Zeitschrift*, 1919, Vol. 3, pp. 314-316.
- [10] L. L. Dines, "On the mapping of quadratic forms," *Bulletin of the AMS*, 47, 494-498 (1941).
- [11] L. Brickman, "On the field of values of a matrix," *Proceedings of the AMS*, 12, 61-66 (1961).
- [12] A. Barvinok, "A course in Convexity," *Graduate Studies in Mathematics*, V. 54, American Mathematical Society, Providence, Rhode Island, 2002.



- [13] F. Bullo, J. Cortes, A. D. Lewis, S. Martinez, “Vector-valued Quadratic Forms in Control Theory,” In V. Blondel and A. Megretski, *Unsolved Problems in Mathematical Systems and Control Theory*, pages 315-320. Princeton University Press, Princeton, NJ, 2004.
- [14] Y. T. Poon, “Generalized numerical ranges, joint positive definiteness and multiple eigenvalues,” *Proc. Amer. Math. Soc.*, 1997, Vol. 125, N. 6, pp. 1625-1634.